Topics for Today

• Network Layer
  – Introduction
  – Addressing
  – Address Resolution

• Readings
  – Sections 5.1, 5.6.1-5.6.2
Network Layer: Introduction

• A network-wide concern!
• Transport layer
  – Between two end hosts
• Data link layer
  – Between two physically connected hosts
• Network layer
  – Involves every router, host in the network
Network Layer Functions

• Addressing
  – **Globally unique** address for each routable device
    • Logical address, unlike MAC address
  – Assigned by network operator
    • Need to map to MAC address

• Forwarding
  – From input port to **appropriate output port** in a router

• Routing
  – **Which path** to use to forward packets from src to dest
Logical View of Router Architecture
Router Input/Output Processing

Input Processing
- Line termination
- Data link processing (protocol decapsulation)
- Look up forwarding, queueing

Switching
- Switching fabric

Output Processing
- Queueing, buffer management
- Data link processing (protocol encapsulation)

Physical layer → Data link layer → Network layer → Switching fabric → Network layer → Data link layer → Physical layer

Data Packet
- MAC addresses
- Network layer addresses
- Network layer payload (data)

Data link layer header
Network layer header
Internet Protocol (IP)

• Universal service in a heterogeneous world
  – IP over everything
• Virtual overlay network
• Globally unique logical address for a host
• Address resolution
  – logical to physical address mapping
IP Addressing

- A 32-bit number that uniquely identifies a location
- Written using dotted decimal notation
- Two-level hierarchy: network id and host id
  - Network IDs administered by
    • Internet Assigned Number Authority
  - Host IDs administered locally
IP Addressing

• IP address is assigned to each network interface
• Routers connect two or more physical networks
  – Each interface has its own address
• Multi-homed host
  – A host having multiple connections to Internet
  – Multiple addresses identify the same host
  – Does not forward packets between its interfaces
IP “Classful” Addressing Scheme

- Three unicast address classes: A, B, and C
- One multicast: class D

```
class
A  0 network  host  1.0.0.0 to 127.255.255.255
B   10 network host  128.0.0.0 to 191.255.255.255
C   110 network host  192.0.0.0 to 223.255.255.255
D   1110 multicast address  224.0.0.0 to 239.255.255.255
```

32 bits
Classless Inter-Domain Routing

- Classful addressing scheme wasteful
  - IP address space exhaustion
  - A class B net allocated enough for 65K hosts
    - Even if only 2K hosts in that network
- Solution: CIDR
  - Eliminate class distinction
    - No A,B,C
  - Keep multicast class D
Classless Addressing

- Addresses allocated in contiguous blocks
  - Number of addresses assigned always power of 2
- Network portion of address is of arbitrary length
- Address format: a.b.c.d/x
  - x is number of bits in network portion of address

```
11001000 00010111 00010000 00000000
```

200.23.16.0/23
IP Addressing

first 24 bits are network address
IP Addressing

Interconnected system consisting of six networks
Special IP Addresses

• Network address: host id = all 0’s
• Directed broadcast address: host id = all 1’s
• Local broadcast address: all 1’s
• Local host address (this computer): all 0’s
• Loopback address
  – network id = 127, any host id (e.g. 127.0.0.1)
Components of a Generic Router
Per Packet Processing in an IP router

1. Accept packet arriving on an incoming link.
2. Lookup packet destination address in the forwarding table, to identify outgoing port(s).
3. Manipulate packet header: e.g., decrement TTL, update header checksum.
4. Send packet to the outgoing port(s).
5. Buffer packet in the queue.
6. Transmit packet onto outgoing link.
Another View of an IP Router

Control Plane

Datapath per-packet processing
Basic Architectural Components

Datapath: per-packet processing

1. Ingress
   - Forwarding Table
   - Forwarding Decision

2. Interconnect

3. Egress
Forwarding Engine

Packet

payload

header

Router

Routing Lookup

Data Structure

Destination

Address

Outgoing

Port

Forwarding Table

<table>
<thead>
<tr>
<th>Dest-network</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>65.0.0.0/8</td>
<td>3</td>
</tr>
<tr>
<td>128.9.0.0/16</td>
<td>1</td>
</tr>
<tr>
<td>149.12.0.0/19</td>
<td>7</td>
</tr>
</tbody>
</table>
Routing Table Lookup is Not an Easy Task

• Search operation is not an exact match
  – Direct lookup needs 4G entries (32 bits IP address)
  – Longest prefix match
  • Hashing table
  • Balanced tree
Size of the Forwarding Table

Source: http://www.telstra.net/ops/bgptable.html
### Lookup Rate Required

<table>
<thead>
<tr>
<th>Year</th>
<th>Line</th>
<th>Line-rate (Gbps)</th>
<th>40B packets (Mpps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998-99</td>
<td>OC12c</td>
<td>0.622</td>
<td>1.94</td>
</tr>
<tr>
<td>1999-00</td>
<td>OC48c</td>
<td>2.5</td>
<td>7.81</td>
</tr>
<tr>
<td>2000-01</td>
<td>OC192c</td>
<td>10.0</td>
<td>31.25</td>
</tr>
<tr>
<td>2002-03</td>
<td>OC768c</td>
<td>40.0</td>
<td>125</td>
</tr>
</tbody>
</table>

$$31.25 \text{ Mpps} \Rightarrow 33 \text{ ns}$$
Basic Architectural Components

Control Plane

Datapath“ per-packet processing
Basic Architectural Components

Datapath: per-packet processing

1. Ingress
   - Classifier Table
   - Forwarding Table
   - Policing & Access Control
   - Forwarding Decision

2. Interconnect
   - Limitation: Memory b/w
   - Limitation: Interconnect b/w
   - Power & Arbitration

3. Egress
   - Limitation: Memory b/w
First Generation Routers

Fixed length “DMA” blocks or cells. Reassembled on egress linecard

Fixed length cells or variable length packets

Typically <0.5Gb/s aggregate capacity
First Generation Routers

Queueing Structure: Shared Memory

Numerous work has proven and made possible:
- Fairness
- Delay Guarantees
- Delay Variation Control
- Loss Guarantees
- Statistical Guarantees

Large, single dynamically allocated memory buffer:
N writes per “cell” time
N reads per “cell” time.
Limited by memory bandwidth.
Second Generation Routers

Typically <5Gb/s aggregate capacity
Second Generation Routers

As caching became ineffective
Second Generation Routers

Queueing Structure: Combined Input and Output Queueing

- 1 write per “cell” time
- Rate of writes/reads determined by bus speed
- 1 read per “cell” time
Third Generation Routers

Switched Backplane

Typically <50Gb/s aggregate capacity
Third Generation Routers

Queueing Structure

1 write per “cell” time

Rate of writes/reads determined by switch fabric speedup

1 read per “cell” time
Third Generation Routers

Queueing Structure

1 write per “cell” time

Rate of writes/reads determined by switch fabric speedup

1 read per “cell” time

Switch

Arbiter

Flow-control backpressure

Per-flow/class or per-output queues (VOQs)

Per-flow/class or per-input queues
Fourth Generation Routers/Switches

Switch Core

Optical links

100's of feet

Linecards
Physically Separating Switch Core and Linecards

- Distributes power over multiple racks.
- Allows all buffering to be placed on the linecard:
  - Reduces power.
  - Places complex scheduling, buffer mgmt, drop policy etc. on linecard.
Fourth Generation Routers/Switches

Switch Core

Optical links

The LCS Protocol

100's of feet

Linecards
Fourth Generation Routers/Switches

Queueing Structure

1 write per “cell” time

1 read per “cell” time

Rate of writes/reads determined by switch fabric speedup

Virtual Output Queues

Switch Fabric

Switch Arbitration

Switch Core (Bufferless)

Lookup & Drop Policy

Lookup & Drop Policy

Lookup & Drop Policy

Output Scheduling

Output Scheduling

Output Scheduling

Typically <5Tb/s aggregate capacity
Problems Facing Router Builders

- Fast forwarding (route lookup)
- Flow identification
  - Fast forwarding path
- QoS support
  - Core routers, enterprise routers, access routers
  - Simpler is winner
- Reliability (stability) of high-speed routers
  - Dual power
  - Duplicate data paths
- Ease of configuration
  - Misconfiguration is another major problem
- Accountability